

2020-08

Affordance matching predictively shapes the perceptual representation of others' ongoing actions.

McDonough, K

<http://hdl.handle.net/10026.1/15481>

10.1037/xhp0000745

Journal of Experimental Psychology: Human Perception and Performance

American Psychological Association

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

**Affordance matching predictively shapes the perceptual representation of others’
ongoing actions.**

Katrina L. McDonough^{a*}, Marcello Costantini^b, Matthew Hudson^{a1}, Eleanor Ward^a, Patric
Bach^a.

^a Plymouth University, School of Psychology, Plymouth, PL48AA, UK

^b Department of Psychological, Health, and Territorial Sciences, “G. d’Annunzio” University
of Chieti-Pescara, Italy and Institute for Advanced Biomedical Technologies - ITAB,
University “G. d’Annunzio”, Chieti, Italy

*Corresponding author: Katrina L. McDonough

Present address:

LSB 0.27, School of Psychology, University of East Anglia, Norwich Research Park,
Norwich, UK, NR4 7TJ

Email: katrina.mcdonough@uea.ac.uk

Word count: 8425 words

¹ Present Address: School of Business, National College of Ireland, Mayor Street, Dublin 1, Ireland.

Affordance matching predictively shapes the perceptual representation of others' ongoing actions.

Abstract

Predictive processing accounts of social perception argue that action observation is a predictive process, in which inferences about others goals are constantly tested against the perceptual input, inducing a subtle perceptual confirmation bias that distorts observed action kinematics towards the inferred goals. Here we test whether such biases are induced even when goals are not explicitly given but have to be derived from the unfolding action kinematics. In two experiments, participants briefly saw an actor reach ambiguously towards a large object and a small object, with either a whole-hand power grip or an index-finger and thumb precision grip. During its course, the hand suddenly disappeared, and participants reported its last seen position on a touch-screen. As predicted, judgments were consistently biased towards apparent action targets, such that power grips were perceived closer to large objects and precision grips closer to small objects, even if the reach kinematics were identical. Strikingly, these biases were independent of participants' explicit goal judgments. They were of equal size when action goals had to be explicitly derived in each trial (Experiment 1) or not (Experiment 2) and, across trials and across participants, explicit judgments and perceptual biases were uncorrelated. This provides evidence, for the first time, that people make on-line adjustments of observed actions based on the match between hand grip and object goals, distorting their perceptual representation towards implied goals. These distortions may not reflect high-level goal assumptions, but emerge from relatively low-level processing of kinematic features within the perceptual system.

Keywords: action understanding; action prediction; social perception; predictive processing; representational momentum.

Public significance statement

- This study shows that our perception of others' actions are biased by which object we expect the person to reach for.
- Hands that form a large grip are seen to reach closer to large objects than they really were, and hands that form a small grip are seen to reach closer to small objects.
- These action expectations are generated during the ongoing observation of other's actions, as grip information becomes available, and is matched to the size of objects within reaching distance.
- These expectations distort the perception of the action, and are independent from explicit action evaluations, suggesting that they emerge from processes within the perceptual system itself, rather than from more sophisticated, high-level reasoning about the actor's goals.

**Affordance matching predictively shapes the perceptual representation
of others' ongoing actions.**

The ability to understand and predict other people's behaviour is a cornerstone of human social cognition and makes people's sophisticated interactions with others possible. Parents rapidly intervene when their child reaches for a hot cup of coffee instead of a toy. In sports, players anticipate each other's behaviour, fluently passing a ball to a team mate's future position. In contrast, deficits in the ability to understand others' behaviour are a hallmark of several conditions that bring with them marked impairments in social interactions, such as autism (e.g., Koster-Hale & Saxe, 2013; Pellicano & Burr, 2012; von der Lühе et al., 2016).

These abilities for social perception are conventionally conceptualized as a simple bottom-up process, in which incoming visual information about others' behaviour is matched to one's higher-level motor – or conceptual – knowledge about it, so that the action's meaning and associated mental states can be derived (Gallese & Goldman, 1998; Rizzolatti & Sinigaglia, 2010; Rizzolatti, Cattaneo, Fabbri-Destro & Rozzi, 2014). However, there is no one-to-one mapping between stimuli and meaning that such a mechanism could rely on (Ansuini, Cavallo, Bertone, & Becchio, 2015; Jacob & Jeannerod, 2005; Uithol, van Rooij, Bekkering, & Haselager, 2011). The same behaviour can mean multiple things in different contexts (e.g. a smile), and the same goals can be accomplished by multiple behaviours (e.g., closing a drawer with one's hand vs hip).

It has therefore been argued that social perception is better understood as a *predictive* process in which the brain constantly tests hypotheses about the observed action against the perceptual input (e.g., Bach, Nicholson, & Hudson, 2014; Bach & Schenke, 2017; Csibra, 2008; Donnarumma, Costantini, Ambrosini, Friston, & Pezzulo, 2017; Kilner, Friston, & Frith, 2007a, 2007b). In such accounts, any assumption about others' goals and beliefs –

derived perhaps from prior knowledge about the individual (e.g., Joyce, Schenke, Bayliss, & Bach, 2015; Schenke, Wyer, & Bach, 2016) or from contextual information (e.g., objects, Bach, Knoblich, Gunter, Friederici, & Prinz, 2005; Jacquet et al, 2012; Kalénine et al., 2014, Nicholson, Roser, & Bach, 2017; gaze and emotional expressions, Adams, Ambady, Macrae, & Kleck, 2006; Frischen & Tipper, 2006) – is translated into predictions about which behaviour should be observed if these assumptions were correct and is superimposed over the perceptual input. Such an integration would not only help to stabilize perception, filling in gaps in the input (e.g. in the case of occlusion, Prinz & Rapinett, 2008) or compensate for the considerable noise during motion perception (Hammett, 1997), but would also let mismatching behaviour of others' stand out, so that our assumptions about them can be revised until they better explain their behaviour.

We recently developed an experimental paradigm that can make these predictions visible (Hudson, Bach, & Nicholson, 2017; Hudson, McDonough, Edwards, & Bach, 2018; Hudson, Nicholson, Ellis, & Bach, 2016; Hudson, Nicholson, Simpson, Ellis, & Bach, 2016; McDonough, Hudson, & Bach, 2019). This paradigm rests on the assumption that, if predictions indeed act on perceptual representations (Bar, 2004; de Lange, Heilbron, & Kok, 2018; Ekman, Kok, & de Lange, 2017) then every prediction one makes about another person may subtly bias the perception of this person's forthcoming actions, especially in the case of uncertainty, such as the visual blurring during motion perception. Thus, in the same way as prior expectations in the non-social world cause us to see a colour differently (Bloj, Kersten, & Hurlbert, 1999; see for an application to the blue/gold dress illusions, Chetverikov & Ivanchei, 2016) or shapes as either convex or concave based on the surrounding illumination (Adams, Graf, & Ernst, 2004), our prior knowledge of other people – their goals and intentions – may subtly shape the perceptual experience of their actions, and help us plan our

own response towards it (see Brouwer & Knill, 2007; 2009 for an example in non-social perception).

This is indeed what we observed in a series of studies. Participants heard an actor make a statement about his goal – “I’ll take it” or “I’ll leave it” – before they briefly saw him start to either reach for an object or withdraw from it, irrespective of what the actor said. The action disappeared mid-motion and participants indicated the perceived vanishing point, either by comparing it to a probe stimulus shortly after stimulus offset (Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016) or by indicating its disappearance point on a touch screen (Hudson et al., 2017; Hudson et al., 2018). The results consistently revealed predictive biases on these perceptual judgments. First, hands were generally reported to have disappeared further along the trajectory than what was actually seen (i.e. closer towards the object for reaches, further away for withdrawals), capturing predictions derived from extrapolating the action’s prior kinematics (i.e. the classical representational momentum effect, Freyd & Finke, 1984; Hubbard, 2005). Second, and more importantly, they revealed an influence of goals attributed to the actor: hands were reported to have disappeared further towards the object when the actor said they would take it and further away from the object when the actor said they wanted to leave it.

Since then, other studies extended these findings, showing that similar distortions can be induced when the participant instructed the observed actor (e.g., telling the actor to “take it!” or “leave it!” before action onset, Hudson, Nicholson, Ellis, et al., 2016) and that the observed actor’s long-term reliability to do as they said modulates the strength of the prediction effects (Hudson et al., 2017). Most recently, we showed that similar effects can be elicited by the prior object context, such that hands reaching straight for an object are perceptually judged to veer slightly upwards if they would need to reach over an obstacle,

and slightly downwards when reaching unnecessarily high (Hudson et al., 2018). Strikingly, the same biases are not observed for inanimate objects (e.g., a ball) that move on closely matched trajectories (McDonough, Hudson & Bach, 2019), linking the perceptual biases to attributions of agency and intentionality to the observed stimulus.

These data show that the goals attributed to others are indeed translated into predictions about their upcoming action, which then bias perceptual judgments towards these expectations. Yet, in all these studies the goals or environmental constraints were explicitly given prior to action onset. In the real world, people typically do not always announce their intentions before their action. Instead, an action's goal often has to be derived dynamically once the action is underway and its kinematics become apparent (Ambrosini, Costantini, & Sinigaglia, 2011; Ambrosini et al., 2013; Bach, Bayliss, & Tipper, 2011; Sartori, Becchio, & Castiello, 2011).

We have argued that the affordances of the goal objects could play a major role in this process (Bach et al., 2014). For example, it is well-known that during natural reaching, the hand forms (“pre-shapes”) a grip that matches the intended target object, long before the object is reached. Viewing an action (e.g. a reach with small grip aperture) that matches – in terms of trajectory, kinematics and hand grip, for example – the affordances of an available target object (e.g. a small object like a strawberry or grape) would immediately signal to an observer what the goal of the action would be (e.g., Bach et al., 2005; Bach, Bayliss & Tipper, 2010; van Elk, van Schie, & Bekkering, 2014). And indeed, there is now ample evidence that people spontaneously derive the target of a reach, by matching the hand's grip configuration – i.e. either a small “precision” grip or a large “power” grip – to the available large or small objects in the environment (e.g., Ambrosini et al., 2011; 2013; for a review, see Bach et al., 2014). For example, eye movements reveal that people anticipate the target of an ongoing reach by matching the unfolding grip shape (large or small grip) to the surrounding

objects before the action is completed (e.g., Ambrosini et al., 2011; 2013), automatic imitation effects are larger for actions that match a goal object (e.g., Bach et al., 2011) and larger motor evoked potential are elicited by the same kinematics if they match an available goal object (Southgate, Johnson, Karoui, & Csibra, 2010).

Here we test, for the first time, whether people use such grip-object matching not only to derive the action's goal or target (e.g., which goal object is selected), but whether they also use these goals, as assumed by perceptual prediction models (e.g., Kilner et al., 2007a, 2007b), to predict how an action is assumed to develop, even if it is already well underway. If this is the case, we should find that the match of an unfolding hand grip to one of two objects in the environment should again induce such perceptual biases, and they should be measureable – as in our prior work – in subtle distortions in perceptual judgment about these actions. Demonstrating such distortions is crucial to show that, during action observation, people go beyond simple goal inference (e.g. identifying the target of a reach) but that they use this information to predict which future course the action will take to achieve these goals, which has only rarely been tested in prior research.

In two experiments, we presented participants with brief videos of an actor's hand starting at rest on the right side of a table, with two potential goal objects – one large, one small – next to each other on the left, one closer in the foreground and the other further into the background. The hand then begins to reach towards the two potential target objects, with the hand forming either a whole-hand power grip or a precision grip, therefore matching one of the two objects but not the other. The hand disappeared mid-motion, at a roughly equal distance – in 3d space – away from either object, and participants were required to indicate the final location of the hand's index finger on a touch screen monitor. If observers identify the goals of the action by matching the observed grip to the two objects' affordances and then

form a perceptual prediction about its future course, then perceptual judgments should show specific biases: the located disappearance points should be reported closer to the corresponding object than they actually were, and away from the alternative (mismatching) target object. Therefore, although the hand actually reached with the same trajectory between the two objects, reaches with a precision grip should be reported closer to the smaller object and reaches with a power grip should be perceptually biased towards the larger object.

A crucial question is whether any such effects emerge from a general top-down mechanism, such that high-level attributions of others' goals penetrate lower-level perceptual representations, or whether any perceptual biases emerge from "encapsulated" interactions in the perceptual system itself (Firestone & Scholl, 2016; Scholl & Gao, 2013), which has already been shown to be sensitive to such matching hand-object interactions (e.g., Bracci & Peelen, 2013). We assumed that the perceptual biases in our prior research (Hudson, Nicholson, Ellis & Bach, 2016, Hudson, Nicholson, Simpson, et al., 2016; Hudson et al., 2018) emerged from high-level information, when people either heard the person make goal statements ("I'll take that!") or instructed them about the appropriate action ("Take it!"). However, in these studies, these goals were given well in advance of the action commencing, so that participants had ample time to "tune" lower-level processes towards the expected goal. To test whether such high-level information can penetrate online social perception – when the action's goal only becomes apparent while the action is underway – we manipulated, across the two experiments, whether participants had to explicitly derive the action's goals or were given no such instruction. In Experiment 1, we had participants say into the microphone, after each trial, which object they thought the hand was reaching for, therefore requiring explicit goal monitoring in each trial. In Experiment 2, no such verbal responses had to be given but participants still reported the perceived disappearance points.

These localisation judgments therefore measure spontaneous goal inferences and the resulting predictions.

These two experiments allow us (1) to replicate any observed biases towards the implied goal object in two separate samples, and (2) to tease apart from which kind of mechanisms these biases emerge. If perceptual biases emerge from a more or less encapsulated, automatic visual prediction system that relies on perceptually available “local” stimulus features (e.g., precision vs. power grip, available objects, for a review, see Scholl & Gao, 2013) then they should be observed – and of roughly equal size – both when the instruction asks participants to view the actions in a teleological manner (Experiment 1) or not (Experiment 2). In contrast, if these biases require attentional focus on the action’s goals, then they should be larger – or only observed – in Experiment 1 where such a focus was induced. Moreover, the combination of explicit verbal goal judgments and implicit perceptual judgments in Experiment 1 will allow us to test, across participants and across trials, the relationship between these measures.

Method

Participants

Sixty-two participants took part in Experiment 1 (mean age = 20 years, SD = 3.4, 52 females) and 63 participants took part in Experiment 2 (mean age = 21 years, SD = 5.5, 50 females). Eleven additional participants across both experiments were excluded due to performance assessed against several a priori criteria (see Results) established in our prior research (Hudson et al. 2018; McDonough et al., 2019). All were right handed and had normal/corrected-to-normal vision, and were recruited from Plymouth University for course

credit. The study was approved by the University of Plymouth Ethics Committee, in accordance with the declaration of Helsinki.

Power analyses were conducted with G*Power (Version 3.1) using the sensitivity analysis function. We do not report theoretical power based on previously reported effect sizes as this neglects uncertainty around these effect size measurements (e.g., Anderson, Kelley & Maxwell, 2017). Instead, we report effect sizes that can in principle be detected with our experimental parameters (i.e. given required power, participant numbers and type of test). This analysis revealed that a sample size of 62 provides .80 power to detect effects in the predicted direction with Cohen's $d = .32$, and effects in either direction with Cohen's $d = .36$. Our prior studies investigating similar effects with the same method (Hudson et al., 2017; Hudson et al., 2018; McDonough et al., 2019) revealed that effect sizes are consistently larger ($d = .52$ to $d = 1.23$). Effect sizes are expected to be smaller here, as, contrary to these prior studies, predictions could only be formed when the action was underway and the prediction-relevant stimulus features were not mentioned to participants before the experiment started. Indeed, a small-scale pilot study with the same design as Experiment 1 showed that the effects would be present, but with a smaller effect size ($d = .39$).

Apparatus

Presentation (NeuroBS) software was used to present the experiment via a HP EliteDisplay S230tm 23-inch widescreen (1920 x 1080) Touch Monitor. Verbal responses for Experiment 1 were detected using Presentation's Sphinx speech recognition engine via a Microsoft LifeChat LX-3000 Headset.

Stimuli

Example stimuli can be seen in Figure 1A. Stimuli were derived, using photo manipulation, from a prior stimulus set of video stimuli from one of the authors (Costantini, Ambrosini, & Sinigaglia, 2012). The videos (950x540) showed an actor's arm, from a slightly elevated side view, reaching over a black table from the right towards two potential target objects adjacent to each other in depth on the left side of the table: a small target (a strawberry) and a large target (an apple). They were derived from videos of natural (e.g. non-pantomimed) reaches towards one of the two objects, which were then photo-edited such that both objects were located at an equal distance (in 3D space) away from the hand, with one closer to the foreground and therefore appearing lower down on the screen and one closer to the background and higher up on the screen (object positions counterbalanced across trials). The two objects were chosen because both their size and typical use clearly affords small “precision” grips (strawberry) and large whole-hand “power” grips (the apple), but they have similar abstract round shapes and merely differ in colour, so that they represent prototypical objects. Both objects cast a small shadow on the surface underneath, re-enforcing the impression that they were resting on the table and were separated in depth.

The actor's hand started at rest in a neutral closed hand posture, and then began to reach, with the hand progressively opening to form either a whole-hand “power” grip or a thumb-index “precision” grip. Four reach videos were used for each hand pre-shape condition (i.e. precision vs. power grip), which together with two target layouts (small object to the front, large at the back, or vice versa), created a total of 16 different videos. Each video was converted into 9 frames, where frames 8 and 9 showed the actor's hand at the approximate point of maximal grip aperture for both the power and precision grip sequences, roughly halfway between the starting position and the objects, and at a position in depth between the

two objects, so that only whether a precision or power grip was formed would predict which object would be reached for and not the hand position. The shadow of the hand was digitally removed so this information could not aid localisation in depth.

Response images for both experiments were created by digitally removing the actor's arm and hand from the scene, so that only the target objects and the background remained.

Presenting this frame immediately after the action sequence gave the impression of the hand disappearing from the scene. A second response image for Experiment 1 was identical to these images, with the addition of four question marks positioned at each corner of the screen. These served as cues for participants to make their verbal responses about which object they believed was the target. All editing was completed using Adobe CC Photoshop.

Procedure

An example trial sequence can be seen in Figure 1B. Participants completed a total of 192 trials, consisting of four blocks of 48 trials (each representing all 16 different trials three times), with breaks in between. At the start of each trial, participants saw an instruction to "Hold the spacebar", to which they pressed the spacebar with their right hand and kept it depressed until the end of the action sequence. They were instructed to keep their left hand in their lap at all times, and this was constantly monitored throughout the experiment. This ensured that they did not track the observed action with their finger, and could only initiate their response with the right hand once the action sequence had disappeared. They then saw the first (neutral) frame of the action sequence for 1000ms, followed by successive frames at 80ms intervals. To increase variability of the hand's final position, the final frame was randomly chosen as either frame 8 or frame 9, such that the hand's intended grip posture was fully visible. This final frame was then immediately replaced with the response image.

Participants released the spacebar and, with their right hand, touched the screen where they thought the final position of the tip of the observed index finger was (see Hudson et al., 2017; Ashida, 2004; Motes, Hubbard, Courtney & Rypma, 2008, for similar procedures). In Experiment 2, the next trial began as soon as the touch response was registered. In Experiment 1, the touch response was immediately followed by the second response frame where participants were required to say into the microphone which target object they thought the actor was reaching towards (either “apple” or “strawberry”). Once the verbal response was registered, the next trial began. The placement of the explicit responses at the end of the action sequences, after perceptual judgements, ensures that explicit verbal responses do not interrupt perceptual judgements, and that the methodology across Experiments 1 and 2 match as closely as possible and can be directly compared.

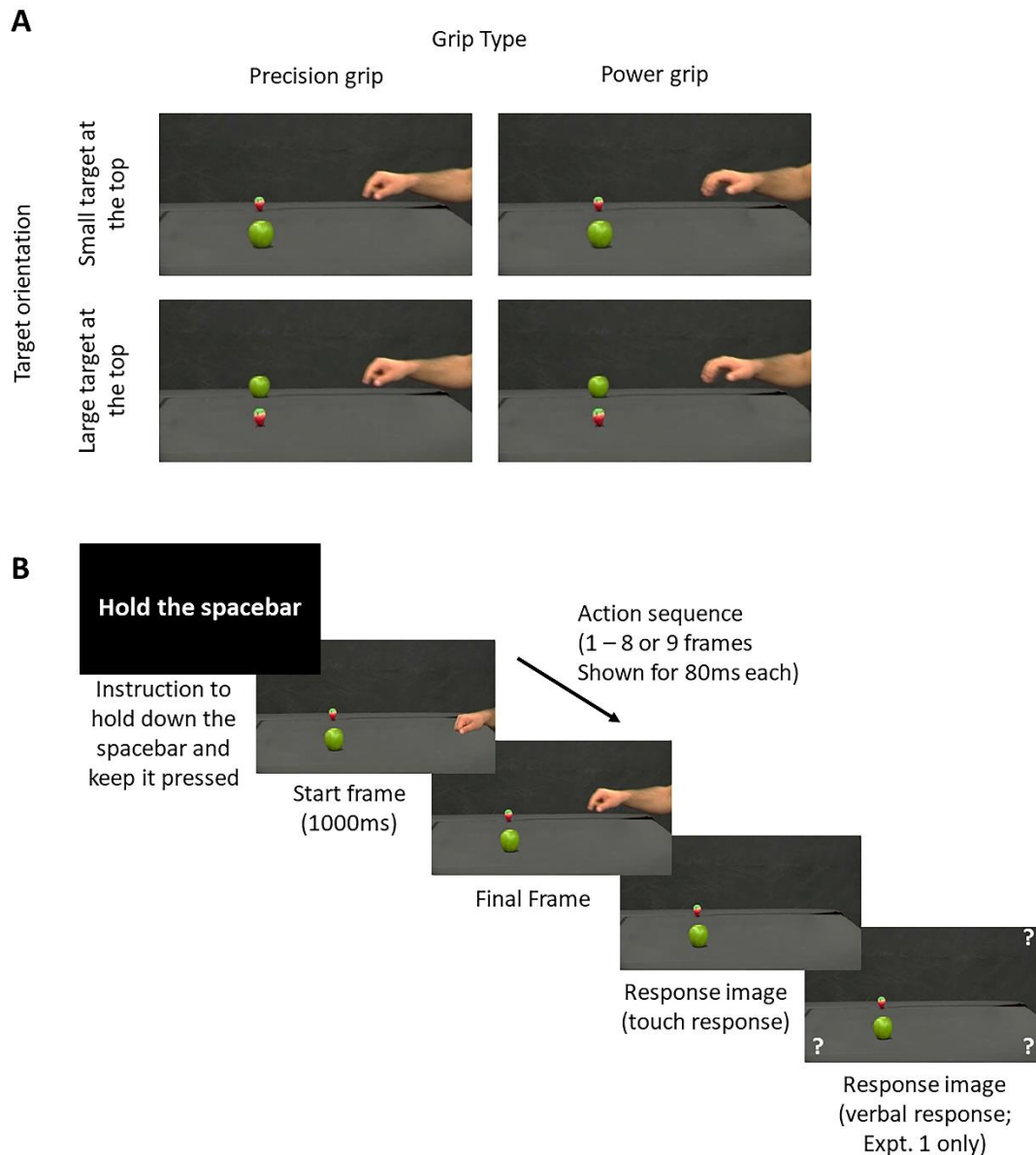


Figure 1. Panel A. Experimental conditions and trial sequence. The objects were arranged with either the small object (strawberry) on top and the large object (apple) on the bottom (top row), or with the large object on top and the small object on the bottom (bottom row). The actor's hand reached with either a precision (small) grip (left column), or with a power (large) grip (right column). Panel B. Example of a trial sequence, showing a “small object on top” configuration with a small grip.

Results

Data were collected in two separate experiments, but for brevity and due to their methodological similarity, we analyse them together in one omnibus ANOVA, with experiment entered as a between-subjects variable. Further separate analyses are presented for each experiment to demonstrate the robustness of the effects.

Data pre-processing

Data filtering was identical to Hudson et al. (2018) and McDonough et al. (2019). In both experiments, individual trials were excluded if the correct response procedure was not followed (e.g. lifting the spacebar before the response image was presented; 2.8% of total trials), or if response initiation or execution times were less than 200ms or more than 3SDs above the sample mean (2.4%, Initiation: mean =355.5ms, $SD=143.7$; Execution: mean =646.0ms, $SD=240.0$). Participants were excluded if too few trials remained after trial exclusions (<50% trials, 5 participants), if their average distance between the real and selected positions was more than 3SDs away from the sample mean (mean =39.0 pixels, $SD=17.0$, 2 participants excluded), or if the correlation between the real and selected positions was more than 3SD below the median r value (X axis: median $r =.762$, $SD = .113$; Y axis: median $r =.860$, $SD = .098$, 2 participants excluded). Two further participants were excluded from Experiment 1, one because they selected the top object as the most likely target object in *all* trials, compared to the remaining participants (top object selected in 51% of trials), and one for showing an abnormally large perceptual bias in the predicted direction (i.e. 15 times larger than the sample mean) so that we suspected a misunderstanding of the task (e.g. touching the likely target object instead of the hand disappearance point). Removal

of these two participants does not affect the results. This left a total of 62 participants in Experiment 1 and 63 participants in Experiment 2.

Perceptual biases

Analysis was conducted on the perceptual bias, which reflects the difference between the hand's real disappearance point and participants' subjective judgments on the X and Y axis in screen coordinates. As in our prior research that established the procedure (Hudson et al., 2018; McDonough et al., 2019), it was derived by subtracting the real final coordinates of the tip of the index finger from the participant's selected coordinates on each trial (see Figure 2). This resulted in separate difference scores for the X and Y axis, where positive X and Y scores represented a rightward and upward displacement respectively, and negative X and Y scores represented a leftward and downward displacement respectively. A score of 0 on both axes indicated that the participant selected the real final position exactly. The coding of negative and positive values on both axes matches the usual convention of screen coordinates. It therefore allows us to represent deviations in localizations as they would appear on the touch screen, relative to the actual disappearance points, without requiring further transformations¹ (see Figure 2A and 2B).

Y axis

Participants' perceptual biases on the Y axis were analysed with a 2x2x2 mixed ANOVA, with Grip type (power vs precision) and Object location (large target on top vs small target on top) as repeated measures factors and Experiment (1: explicit prediction vs 2: implicit prediction) as between-subjects factor. We predicted, first, that the perceived disappearance

points of reaches would be distorted towards their apparent target object, anticipating the action's predicted future course. As a consequence, reaches should appear to have terminated slightly higher if they matched a target object at the top and lower for a target object at the bottom. Second, if these perceptual biases depend on viewing the actions teleologically, in terms of their goals, then these shifts should be larger, or only observed, in Experiment 1, where the hands' goals were task-relevant, compared to Experiment 2 where such goal inferences would only occur spontaneously. These effects should be reflected in a two-way interaction of Grip type and Object location and a three-way interaction of Grip type, Object location, and Experiment. Note that each participant's contrast value for the two-way interaction of Grip type and Object location is mathematically identical to their overall bias in judgments towards the target object on the Y-axis, as typically reported in studies on representational momentum (O displacement, Hubbard, 2005; Hubbard & Bharucha, 1988). We present means and single data points on this summary measure for each participant in raincloud plots (Allen, Poggiali, Whitaker, Marshall & Kievit, 2018) for both experiments in Figure 2C.

As can be seen in Figure 2 A & B, disappearance judgements were generally lower than the real position, consistent with previously reported shifts of localization responses towards the objects's centre of gravity in touch screen studies (e.g., Coren & Hoenig, 1972; Hudson et al., 2017; McDonough et al., 2019) and a general downward bias in localization responses in representational momentum-like paradigms (i.e. representational gravity, Hubbard, 2005).

The omnibus ANOVA revealed the predicted interaction of Grip type and Object location, $F(1,123) = 16.8, p < .001, \eta_p^2 = .120, BF_{10}=158$. Even though kinematics were identical, the disappearance point of power grips was reported higher when the large target object was placed at the top (-15.1px) than when the large target object was placed at the bottom (-

16.0px, $t(124)=2.75$, $p=.007$, $d=.25$). Conversely, the disappearance point for precision grips was reported to be higher when the small target object was at the top (-1.5px) compared to when the small target object was at the bottom (-2.8px, $t(124)=3.64$, $p<.001$, $d=.33$). Note that while these deviations are small, they are highly reliable, reflected in a Bayes factor of BF10 of 158 for the interaction, and – especially when their smaller than lifelike representation on the monitor is considered – surpass similar deviations induced by object biases in the kinematics of real reaches (e.g., distractor interference/deviation ≈ 1 mm, Keulen, Adam, Fischer, Kuipers & Jolles, 2003; Welsh, Elliott & Weeks, 1999).

The second question was whether the size of these perceptual displacements was larger when the actions' target was task relevant in Experiment 1 compared to when it was task-irrelevant in Experiment 2. However, there was no three-way interaction between grip type, object location and experiment ($F(1,123) = .666$, $p = .416$, $\eta_p^2 = .005$, BF10=.263). Moreover, the relevant interaction of grip and object location was present in both experiments, irrespective of whether participants explicitly reported the action's goals after the perceptual judgments (Experiment 1: $F(1,61) = 9.56$, $p = .003$, $\eta_p^2 = .135$, BF10=30.2, Experiment 2: $F(1,62) = 7.20$, $p = .009$, $\eta_p^2 = .104$, BF10 = 11.0, see Figure 2). Two one-sided tests (TOST) procedure (Lakens, 2017) indicated that the observed effect size of the between-experiment difference ($d=.16$) was significantly within the equivalence bounds of $\Delta L = -.51$ and $\Delta U = .51$, $t(113.89) = -1.95$, $p = .027$, and a Bayesian analysis provides substantial evidence for the null hypothesis, BF10 = .110.

As all other effects in the ANOVA were not predicted, they should be treated as incidental findings, unless they pass a threshold of $p < .01$, corrected for multiple comparisons in an ANOVA (Cramer et al., 2015). Only the main effect of grip type, $F(1,123) = 1130$, $p < .001$, $\eta_p^2 = .902$, BF10=3.20e+141, surpassed this threshold, showing perceived disappearance

points of power grips were displaced further downward than precision grips. This was expected since the power grip is larger and therefore has a lower centre of gravity, which is known to affect touch screen judgements, but is independent from our effects of interest (for similar findings, see Coren & Hoenig, 1972; see also Hudson et al., 2017; Hudson et al., 2018). There were no further main effects or interactions (all $F < 1.62$, all $p > .205$, all $BF_{10} < .472$, see supplementary table for a report of all effects).

X axis

Perceptual biases on the X axis were analysed with the same ANOVA model. As the two target objects were equidistant from the hand disappearance point on the X axis, the apparent matching of grip to the top or bottom object should not affect deviations on the X axis. All effects in this ANOVA are therefore unpredicted and incidental, and subject to alpha inflation due to multiple comparisons in an ANOVA (Cramer et al., 2015). They should therefore be evaluated against a Bonferroni-adjusted alpha of $p < .007$. Only the main effect of grip type, $F(1,123) = 503$, $p < .001$, $\eta_p^2 = .804$, $BF_{10} = 1.17e+99$, passed this adjusted threshold, with the perceived disappearance point of power grips more leftward than precision grips, which again reflects the well-known biases in touch screen responses to the more leftward centre of gravity for the more spatially isolated index finger in power grips than precision grips. There were no further main effects or interactions (all $F < 4.07$, all $p > .046$, all $BF_{10} < .226$, see supplemental material for details).

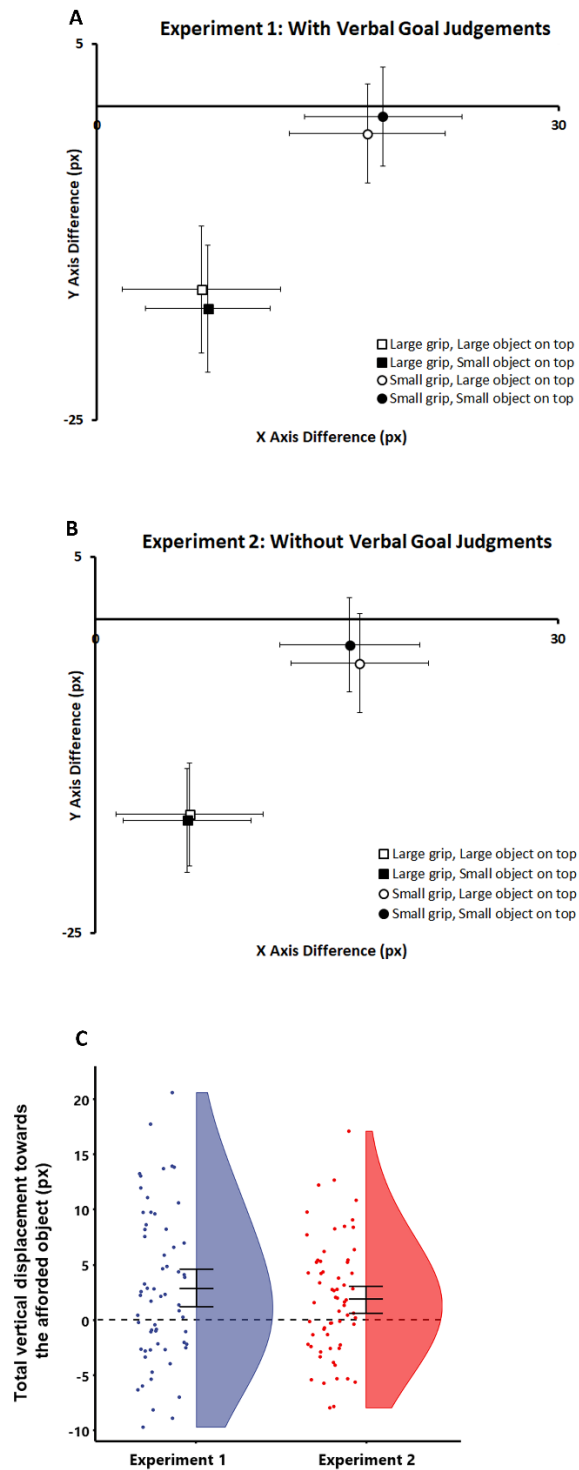


Figure 2. Panel A. Grip type x object interaction for experiment 1. The difference scores between the real final position and the selected final position is plotted for the X axis and Y axis. Panel B. Grip type x object interaction for experiment 2. Panel C. A raincloud plot (Allen, Poggiali, Whitaker, Marshall & Kievit, 2018) of the comparison across experiments of the size of the Y axis interaction in pixels, equivalent to the total amount by which each grip type was distorted towards the congruent object (mathematically equivalent to the O displacement in representational momentum studies, Hubbard, 2005; Hubbard & Bharucha, 1988). Each data point represents this Y axis interaction value for each participant. Error bars in all plots depict 95% confidence intervals.

Relationships between perceptual shifts and explicit goal judgments

An important question is to what extent the perceptual displacements measured above are informed by people's higher-level, explicit judgments about the observed actions. If perceptual predictions are shaped by high-level goal attributions or vice versa, then perceptual displacement and explicit judgments in Experiment 1 should be closely linked, both across participants and across trials within participants.

Do people rely on affordance matching to make explicit goal judgments?

We first established whether grip/object-matching does not only inform perceptual biases in action observation, but also people's *explicit* judgments about the action's goals. Participants were never given explicit instruction about the relevance of grip/object match, but they reported, after each action in Experiment 1, whether they subjectively experienced the hand to be reaching for the apple or the strawberry. To test whether these judgments were informed by the apparent grip/object-match, we separately coded which object the grip *actually* corresponded to and participants' subjective judgments about which object they *felt* the hand reached for. To this end, a hand grip matching the bottom object was coded as 0 and a match to the top object was coded as 1. Verbal goal judgments were similarly coded as 0 and 1 for perceived goal objects at the bottom and at the top, respectively. We then simply, for each participant, calculated the proportion of verbal goal judgments that corresponded to the actual match with the goal object. A simple t-test against chance (50%) revealed that explicit judgments followed the actual hand-object match ($M = 66.7\%$, $SD = 16.7\%$; $t = 31.5$, $p < .001$). This confirms that the grip-object-match did not only inform perceptual

displacements (see main analysis), but also participants' explicit goal object judgments. It is important to note that even in the explicit judgments grip-object matching only induced a subtle shift towards the intended goal object (i.e. 17% away from chance performance). This is consistent with the fact that grip-object matching is only one of many potential stimulus features (e.g. direction, kinematics) in natural reaches that provides information about the action's target, which was not specifically mentioned to participants. The shift in explicit judgments therefore measures participants' spontaneous use of this feature, allowing us to test in how far this explicit use is linked to the measured biases in perceptual judgments, both across participants and across the trials within the experiment.

Are perceptual biases and explicit judgments related across participants?

We then tested whether individual differences in participants' tendency to rely on grip/object matching to make explicit judgments is related to their reliance on grip/object matching in perceptual judgments. We therefore correlated the proportion to which each participant's verbal goal judgments matched the actual hand/object match with their perceptual biases towards the grip-matching goal object (i.e. the interaction contrast that marks the predictive perceptual shift due to matching grips to object affordances in the main analysis above, mathematically identical to the O displacement in representational momentum studies, Hubbard, 2005; Hubbard & Bharucha, 1988). Surprisingly, the two types of judgments were almost perfectly uncorrelated, $r(59) = .08$, $p = .518$, $N=62$. TOST procedure (Lakens, 2017) indicated that the observed effect size ($r=.08$) was significantly within the equivalence bounds of $\Delta L = -.36$ and $\Delta U = .36$, $p = .011$. A Bayesian analysis provided substantial evidence for the absence of a correlation, $BF_{10} = .102$.

Are perceptual biases and explicit judgments related across trials?

While there may be no overall relationships between a participants' perceptual biases and explicit judgment, it is possible that such relationships are present on a trial-by-trial basis. If explicit and perceptual judgments depend on one another, then actions/trials judged explicitly to be directed towards the top object should also be more likely to show a perceptual mislocation towards the top than the bottom, and vice versa for reaches judged to be directed to the bottom object. As these trial-by-trial relationships are independent of the overall (mean) value on both measures, they can be observed irrespective of any across-participants relationship between perceptual and explicit judgments.

To test this, we first verified that the perceptual shifts in each trial reflected the *actual* target object location. We therefore correlated, for each participant separately, the actual target object location (coded as 0 or 1) for each trial with the size of the perceptual judgment displacement on the Y axis across all trials of the participant. Testing the resulting fisher-transformed correlation coefficients against zero with a simple t-test revealed a positive mean correlation between perceptual shifts and target object location across participants (mean $r = .03$, $t = 3.10$, $p = .003$; $d = .40$), see Figure 3A. This trial-by-trial correlation was replicated in Experiment 2 (mean $r = .02$, $t = 2.77$; $p = .007$; $d = .35$), and did not differ from Experiment 1 ($t(123) = .512$, $p = .609$, $d = .14$). TOST procedure (Lakens, 2017) indicated that the observed effect size between the experiments ($d = .14$) was significantly within the equivalence bounds of $\Delta L = -.51$ and $\Delta U = .51$, $t(121.5) = -2.06$, $p = .021$. As before, a Bayesian analysis of this effect provided substantial evidence for an absence of such a difference, $BF_{10} = .134$.

Replicating the results of the main analyses with an across-trials correlational measure, this analysis therefore confirms that actions in which the hand grip matched the top object induced larger shifts upwards than a grip match to the bottom object.

Next, we performed an identical across-trials correlation analysis for the relationship between the actual target object location and verbal goal judgment. Replicating the finding from the main analysis that explicit judgments follow grip-matching, this again revealed a positive correlation between verbal goal judgments and grip information (mean $r = .44$, $t = 7.20$, $p < .001$; $d = .91$), see Figure 3B. Across trials, reaches whose grip matched the top object were therefore more likely to be judged to be reaching to the top object, and vice versa for reaches whose grip matches the bottom object.

Finally, the crucial question was whether explicit verbal goal judgments about an action and the perceptual shifts showed a positive relationship. Strikingly, as in the across-participants analysis, there was no correlation between perceptual displacements and the explicit verbal goal judgments across trials (mean $r = .01$, $t = .694$, $p = .500$, $d = .09$), see Figure 3C. TOST procedure (Lakens, 2017) indicated that the observed effect size ($d = .09$) was significantly within the equivalence bounds of $\Delta L = -.36$ and $\Delta U = .36$, $t(61) = -2.15$, $p = .018$, and a Bayesian analysis provided strong evidence for the absence of a relationship, $BF_{10} = .088$. Thus, while the actual target location directly related to the perceptual judgments and to the verbal goal judgments, the two types of judgments were not related to each other. Figure 3 shows participants' individual across-trial correlation coefficients for the three correlations of interest.

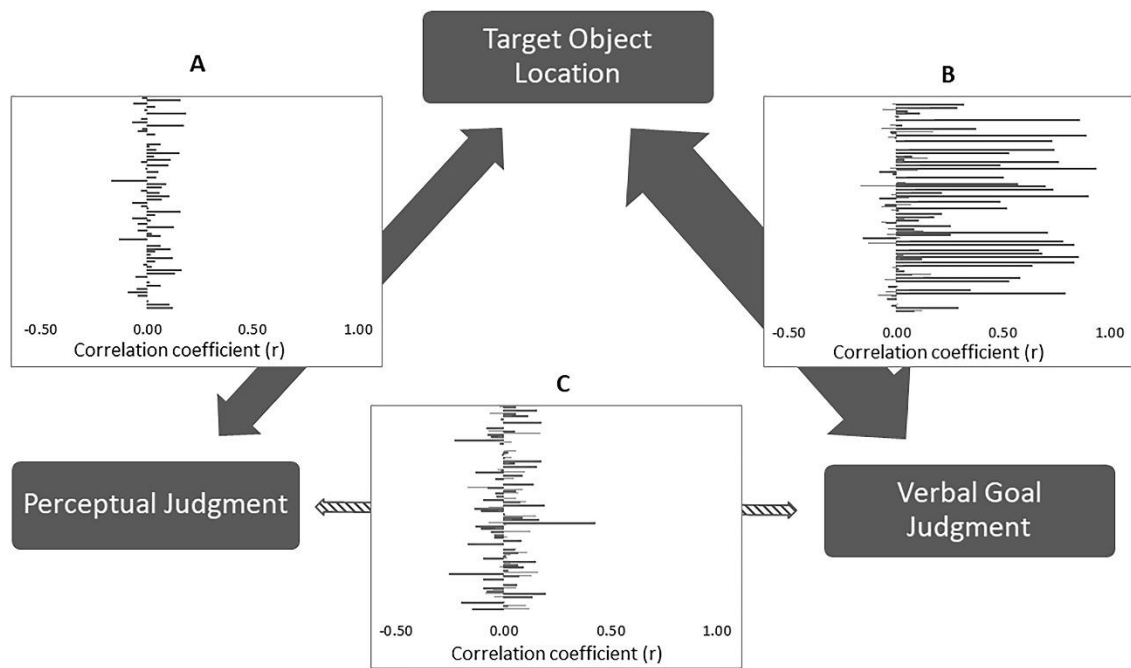


Figure 3. Individual participants' across-trial correlation coefficient (r) between the three measures in interest. The horizontal bars in each of the three panels shows each participant's pairwise across-trial correlation coefficients for the correlation between perceptual judgments and the target object location (top object or bottom object, based on the grip-object match, Panel A), the correlation between verbal goal judgments and the target object location (Panel B), and the correlation between perceptual judgments and verbal goal judgments (Panel C).

Discussion

Prior work has shown that people integrate object and action kinematic information to derive the likely goal of observed actions, even while the action is still ongoing (Ambrosini et al., 2011; Bach et al., 2011; Decroix & Kalénine, 2018; Eshuis, Coventry, & Vulchanova, 2009; for a review, see Bach et al., 2014). Here, we tested the hypothesis of predictive processing models (Bach & Schenke, 2017; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016) that such goal inferences are immediately translated into perceptual predictions about the actions future path towards the inferred goal, and bias perceptual judgments towards these expected trajectories.

The data from two experiments supported this proposal. In each trial, participants observed the initial stages of a reach towards two potential target objects that differed in size, with either a whole-hand power grip or a precision grip, and were asked to localize the hand's last seen position after its sudden offset. The results revealed consistent biases in perceptual judgments towards action expectations derived from the match between the emerging grip type and object size. While reaches with a power grip were reported to be closer to large objects, reaches with a precision grip were perceived to be closer to small objects, even when actions with the same kinematics were observed and only the location of the relevant target object changed. These perceptual mis-locations were present both when participants were explicitly asked to identify the goal objects in a secondary task (Experiment 1) and when the reach targets were completely task irrelevant and participants were only asked to accurately report the hand's disappearance point (Experiment 2). Moreover, they were observed even though the actor's hand started at rest. Which of the two objects was the target object was therefore ambiguous before action onset and only became apparent once the action commenced and a specific grip type began to form.

These perceptual displacements towards the expected kinematics support predictive processing models of social perception, which argue that any inferences about an observed action's goal will (1) give rise to predictions about the action's further kinematics, which can then (2) bias action perception towards these expectations (Bach & Schenke, 2017; Hudson et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016; Kilner et al., 2007a, 2007b). They go beyond previous findings in which action expectations were explicitly induced prior to action onset, for example, by asking participants to instruct the (virtual) actor (Hudson, Nicholson, Ellis, et al., 2016), by hearing goal statements of the actor ("I'll take it!", Hudson et al., 2017; Hudson, Nicholson, Simpson, et al., 2016), or by presenting a static image of the goal object with or without obstructing objects in the way

(Hudson et al., 2018; McDonough et al., 2019). Here, no such prior information was available. The actions started from a neutral position and the goals only became apparent once the action was underway, from the subtle pre-shaping of the hands for the affordances of the goal object (i.e. precision grip when directed towards the small object, power grip when directed towards the large object).

Our results therefore show, first, that predictions are not just made before action onset, but are dynamically adjusted “on-line” as more information becomes available from the unfolding kinematics and are then integrated with the action’s perceptual representation. Second, they reveal that matching of actions to the affordances of potential goal objects in the environment plays a major role in this process, as previously hypothesized (see Bach et al., 2014 for a theoretical proposal and review). Third, they go beyond prior work that has shown that people use such affordance matching to identify the *target* of another’s action, guiding eye movements towards it in an anticipatory manner (e.g., Ambrosini et al., 2011; Ambrosini et al., 2013; Bach et al., 2011). They reveal that these predictions do not just represent the likely goal object, but represent concrete expectations about the next step of a hand’s path through the scene, which are perceptually integrated with the actually observed kinematics.

A surprising finding for such theories of top-down guided predictive perception was that this matching of actions to goal objects and the resulting perceptual biases appeared to be highly automatic and independent from explicit judgments about the action’s goals. We had hypothesized that if perceptual mislocations and explicit judgments inform each other, then those trials that were explicitly judged to be directed towards the top object should also show upwards mis-locations, and vice versa for judgments towards bottom objects. However, converging findings from three independent tests argue against this interpretation. First, in Experiment 1, the action’s goal was highly task relevant because participants indicated

verbally, after each action, which object they believed was the target. No such response was required in Experiment 2, such that any perceptual bias indexes only spontaneous, implicit goal inferences and predictions that are nevertheless made by participants. Nevertheless, the perceptual bias towards the grip-matching object was evident – with virtually identical effect sizes – in both experiments. Second, in Experiment 1, correlational analyses across participants showed that the perceptual biases were independent of whether participants’ verbal goal judgments revealed a reliance on grip information or not. Thus, the perceptual bias towards the matching goal object was of similar size irrespective of whether participants made use of grip-object matching in their explicit post-action judgments. Third and finally, correlational analyses across trials that directly relate perceptual displacements in a given trial to verbal goal judgments in the same trial (Experiment 1) confirmed this lack of a top-down influence. Even though the actual goal object predicted both the direction of the perceptual bias and which object was explicitly reported as a target, the perceptual biases and verbal goal judgments remained uncorrelated across trials.

Together therefore, the results of these three independent tests show that, while the mechanisms for action prediction and the goal identification both rely on grip-object matching, the two mechanisms do not strongly inform each other: explicit goal judgments do not induce perceptual biases, nor do perceptual biases induce explicit goal judgments. While it had been difficult to draw conclusions from null effects, recent Bayesian and equivalence-testing (TOST) analysis techniques that were used here have overcome this problem (Lakens, 2017). Indeed, all three tests described above provide substantial to strong evidence *for* the absence of a link between explicit judgments and perceptual biases, and the reported difference remain significantly within equivalence bounds.

This apparent dissociation between explicit and implicit perceptual biases is in line with prior reports of similar dissociations during motion perception in representational momentum task, where people's abstract knowledge of an object's future course deviates from their perceptual predictions (e.g., Freyd & Jones, 1994; Kozhevnikov & Hegarty 2001). It may appear surprising, however, from a viewpoint of recent predictive coding accounts, according to which predictions and prediction errors ensure that top-level and lower-level judgments remain aligned (Clark, 2013; Friston & Kiebel, 2009). Thus, any inferences on a higher level – for example, what the goal of the action is – would propagate downwards to lower levels and inform perceptual judgments. Conversely, any change in perceptual estimation – whether the hand is perceived to travel upwards and downwards – would, via prediction errors, inform resulting high-level judgments of action goals.

This apparent conflict can be resolved if one accepts recent proposals that predictions can also emerge locally, from top-down interactions *within* the human perceptual system for the perception of biological motion (for a review, see Scholl & Gao, 2013), without drawing on information external to these networks such as high-level explicit action goal judgments (e.g., Firestone & Scholl, 2016). Indeed, several lines of evidence suggest that the perceptual system itself can detect many aspects of intentional behaviour, without the need for higher-level evaluation, such as whether one actor chases another (Gao, McCarthy & Scholl, 2010) or whether actors pay attention to their reach or the target (e.g. Jellema, Baker, Wicker, & Perrett, 2000). The match between hand and goal object may therefore provide another feature from which such lower-level teleological interpretations of observed motion can be derived, and low-level perceptual regions have indeed been found to be sensitive to such information (e.g., STS, Gao, Scholl & McCarthy, 2012; Saxe, Xiao, Kovacs, Perrett & Kanwisher, 2004; lateral occipital cortex, Bracci & Peelen, 2013). Our new data then suggests that these regions are not only sensitive to the presence of these matches but that

they also use them to predict the action's further path and bias the perceptual representations towards it, independently of the goals *explicitly* attributed to the other person.

Future studies now need to investigate from what kind of mechanism the perceptual biases emerge. Our prior studies point towards lower-level perceptual processes that determine participants' conscious perceptual experience of the actions, which then drives their explicit judgments. Several studies from our lab (Hudson et al., 2017; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016) have shown, for example, that the type of perceptual displacements measured here emerge not only with touch screen judgments, which may also have a motor (pointing) component, but also purely perceptual judgments, in which participant compare the hand disappearance points with probes in the same, future, or prior position. Even when these probes were presented only 250 ms after hand disappearance, the same perceptual distortions were apparent, suggesting, at the very least, an effect in iconic memory. Second, in our most recent work testing perceived changes to action kinematics in the presence of obstacles, all such effects were eliminated when dynamic visual noise masks were presented briefly (560 ms) between action offset and touch screen response (Hudson et al., 2018), which are known to interfere with re-entrant top-down projections to early visual cortex (Fahrenfort, Scholte, & Lamme, 2007; Lamme, Zipser, & Spekreijse, 2002).

Together, therefore, these findings support a low-level locus of the effects that either reflects the top-down sharpening of perceptual representations during motion perception (i.e. motion blurring, Hammett, 1997), or the filling-in of the expected path after the unexpected sudden offset (Ekman et al., 2017). Neuroimaging studies would be useful to disentangle to what extent the perceptual changes we have measured here reflect changes in early perceptual systems, similar to that seen in various visual illusions and sometimes in motion illusions

(e.g. apparent motions, Muckli, Kohler, Kriegeskorte, & Singer, 2005; predicted motion pre-play, Ekman et al., 2017).

Another challenge is elucidating whether – and via which mechanisms – the perceptual predictions ultimately influence cognition and behaviour. On first glance, the relatively small absolute size of the effects (~ 1 pixel) may suggest an only negligible influence in real-life social interactions. However, this needs to be evaluated against the background of a not very salient stimulus-feature that is available only late during action observation, and that, even when participants' were explicitly asked, only induced a 17% shift away from chance performance. On this view, it is striking that robust shifts in perceptual experience are observed at all, with an effect size that is comparable to other effects in psychology (Schafer & Schwarz, 2019). Indeed, in our prior work in which the decks were not stacked against finding perceptual biases and the prediction-relevant stimulus features were available earlier and much more salient (e.g. the spatial setup of the scenes, intention statements of the actors), the effects were much larger, between $d=.52$ and $d=1.23$ (e.g., Hudson, Nicholson, Ellis & Bach, 2016, Hudson et al., 2017, McDonough et al., 2019).

The question about how the measured biases would manifest in real-life social interactions is also complicated by the fact that there is still considerable debate about how, specifically, predictions would influence behaviour, with different theoretical frameworks making different predictions. For example, in Bayesian-perception views, predictions primarily serve to “sharpen” ambiguous social perception (e.g., Yon, Gilbert, de Lange & Press, 2018; Kok, Jehee, de Lange, 2012) or to “fill in” missing or ambiguous stimulus aspects (e.g. during occlusion or when saccading away). The influence of such a sharpening depends inversely on the reliability one attributes to the real input. In contrast to here, in real-life situations, participants do not closely watch the actions and report their spatial extent. They see actions

often only briefly before saccading away (e.g. towards the anticipated goal object), they plan their own responses at the same time, and the actions take place in crowded environments with unpredictable onsets. All these factors would conspire to reduce attention to the seen movements, thereby decreasing their reliability and in turn increasing the reliance of top-down predictions beyond what is measured in our experiments.

In other views, predictions do not primarily serve perceptual functions, but to help “bridge the gap” between relatively slow perceptual and quick, online motor control processes (Hubbard, 2006, see also Müsseler, Stork & Kerzel, 2008; Nijhawan, 1994, 2008). In such a view, the weight given to expectations increases the more the observed action needs to provide a control signal for own action control (e.g. intercepting an observed reach), and their effect is best measured not in resulting perceptual distortions but in anticipatory changes to these actions. Indeed, several studies have shown that perceptual distortions can have different – and usually larger – effects when measured directly on the actions directed at these stimuli (e.g., Brouwer & Knill, 2007; 2009), and that perceptual biases themselves increase the more they serve as control signal for the observer’s actions (Kerzel, 2003; Müsseler et al., 2008).

Finally, recent predictive-processing approaches assume that the primary role of predictions is neither to primarily serve perception nor to guide action (even though both are relevant contributions), but to perceptually test one’s prior assumptions about the other person against their actual behaviour, so that, in the long run, one’s knowledge of the (social) world remains aligned with reality (e.g., Clark, 2013; Friston & Kiebel, 2009; see Bach & Schenke, 2017, for an application to social perception). In such views, the subtle biases in visual judgments measured in our paradigm show that such perceptual comparisons are being made, but they

say nothing about their ultimate role in shaping social judgments, when an action clearly mismatches the prior predictions.

Together, therefore, the perceptual distortions measured here and in our prior work show that people make predictions about the course of others' actions, and that these predictions are available in a format that is dynamically integrated with the perceptual input, even when the action is already underway. They may therefore provide a promising starting point in disentangling these, not mutually, exclusive roles that perceptual predictions could play in human interactions with the (social) world.

Conclusions

The present results reveal that the perceptual experience of others' actions is predictively shaped by the integration of the unfolding action kinematics with the affordances of available goal objects, as proposed by recent predictive models of social perception, (Csibra, 2008; Kilner et al., 2007^{ab}; Bach, Nicholson & Hudson, 2014; Schenke & Bach, 2017). These integrations likely emerge at a relatively low-level, from processes within systems for the perception of biological motion, without influences from top-down evaluations of others' goals and intentions. Future studies must now resolve precisely via which mechanism predictions act on perceptual representations, how they help guide own actions towards future states in social interactions, and how prior knowledge is updated and revised if it consistently fails to explain the perceptual input.

Author note

A pre-print of this manuscript has been uploaded to PsyArXiv, available at <https://psyarxiv.com/mbea2>. The data in this manuscript are available at <https://osf.io/des8x/>.

The results have been presented as a poster at the Future of Social Cognition conference (University of East Anglia, UK, 2018), MeeTo (University of Turin, Italy, 2018) and the British Association for Cognitive Neuroscience Annual Meeting (University of Plymouth, UK, 2017).

Authors' contributions

KLM, MC, MH, EW & PB devised the experiment. Stimuli were created by MC and edited by KLM. The program was developed by MH and KLM. Data was collected by KLM. Data were analysed by KLM, EW & PB. The manuscript was written by KLM, PB and MH and EW provided critical feedback. All authors gave final approval for publication.

Funding

This work was supported by the Economic and Social Research Council [grant number ES/J019178/1] awarded to Patric Bach, and a University of Plymouth doctoral student grant awarded to Katrina L. McDonough.

Competing interests

The authors declare no competing interests.

References

- Adams, R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion*, 30(2), 177-186. doi:10.1007/s11031-006-9020-2
- Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience*, 7(10), 1057-1058. doi:10.1038/nn1312
- Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. (2018). Raincloud plots: a multi-platform tool for robust data visualization. *PeerJ Preprints*, 6, e27137v1.
- Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. *Journal of Neurophysiology*, 106(3), 1437-1442. doi:10.1152/jn.00118.2011
- Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking Ahead: Anticipatory Gaze and Motor Ability in Infancy. *PLOS ONE*, 8(7), e67916. doi:10.1371/journal.pone.0067916
- Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychological Science*, 28(11), 1547-1562.
- Ansuini, C., Cavallo, A., Bertone, C., & Becchio, C. (2015). Intentions in the brain: the unveiling of Mister Hyde. *The Neuroscientist*, 21(2), 126-135.
- Ashida, H. (2004). Action-specific extrapolation of target motion in human visual system. *Neuropsychologia*, 42(11), 1515-1524.
- Bach, P., & Schenke, K. C. (2017). Predictive social perception: Towards a unifying framework from action observation to person knowledge. *Social and Personality Psychology Compass*, 11(7), e12312. doi:10.1111/spc3.12312
- Bach, P., Bayliss, A. P., & Tipper, S. P. (2011). The predictive mirror: interactions of mirror and affordance processes during action observation. *Psychonomic Bulletin & Review*, 18(1), 171-176. doi:10.3758/s13423-010-0029-x
- Bach, P., Knoblich, G., Gunter, T. C., Friederici, A. D., & Prinz, W. (2005). Action Comprehension: Deriving Spatial and Functional Relations. *Journal of Experimental Psychology: Human Perception and Performance*, 31(3), 465-479. doi:10.1037/0096-1523.31.3.465
- Bach, P., Nicholson, T., & Hudson, M. (2014). The affordance-matching hypothesis: how objects guide action understanding and prediction. *Frontiers in Human Neuroscience*, 8. doi:10.3389/fnhum.2014.00254
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5(8), 617-629. doi:10.1038/nrn1476
- Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature*, 402(6764), 877-879. doi:10.1038/47245

- Bracci, S., & Peelen, M. V. (2013). Body and Object Effectors: The Organization of Object Representations in High-Level Visual Cortex Reflects Body-Object Interactions. *Journal of Neuroscience*, 33(46), 18247-18258. doi:10.1523/jneurosci.1322-13.2013
- Brouwer, A-M. & Knill, D. C. (2007). The role of memory in visually guided reaching. *Journal of Vision*, 7, 1-12. <https://doi.org/10.1167/7.5.6>
- Brouwer, A-M. & Knill, D. C. (2009). Humans use visual and remembered information about object location to plan pointing movements. *Journal of Vision*, 9, 1-19. <https://doi.org/10.1167/9.1.24>
- Chetverikov, A., & Ivanchei, I. (2016). Seeing “the Dress” in the Right Light: Perceived Colors and Inferred Light Sources. *Perception*, 45(8), 910-930. doi:10.1177/0301006616643664
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(03), 181-204. doi:10.1017/s0140525x12000477
- Coren, S., & Hoenig, P. (1972). Effect of Non-Target Stimuli upon Length of Voluntary Saccades. *Perceptual and Motor Skills*, 34(2), 499-508. doi:10.2466/pms.1972.34.2.499
- Costantini, M., Ambrosini, E., & Sinigaglia, C. (2012). Out of your hand's reach, out of my eyes' reach. *Quarterly Journal of Experimental Psychology*, 65(5), 848-855. doi:10.1080/17470218.2012.679945
- Cramer, A. O. J., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P. P. P., . . . Wagenmakers, E.-J. (2015). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review*, 23(2), 640-647. doi:10.3758/s13423-015-0913-5
- Csibra, G. (2008). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rosetti, & M. Kawato (Eds.), *Sensorymotor foundations of higher cognition. Attention and performance XXII* (pp. 435-459). Oxford: Oxford University Press.
- de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, 22(9), 764-779. doi:10.1016/j.tics.2018.06.002
- Decroix, J., & Kalénine, S. (2018). Timing of grip and goal activation during action perception: a priming study. *Experimental Brain Research*, 236(8), 2411–2426. doi:10.1007/s00221-018-5309-0
- Donnarumma, F., Costantini, M., Ambrosini, E., Friston, K., & Pezzulo, G. (2017). Action perception as hypothesis testing. *Cortex*, 89, 45-60. doi:10.1016/j.cortex.2017.01.016
- Ekman, M., Kok, P., & de Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, 8, 15276. doi:10.1038/ncomms15276
- Eshuis, R., Coventry, K. R., & Vulchanova, M. (2009). Predictive Eye Movements Are Driven by Goals, Not by the Mirror Neuron System. *Psychological Science*, 20(4), 438-440. doi:10.1111/j.1467-9280.2009.02317.x

- Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2007). Masking Disrupts Reentrant Processing in Human Visual Cortex. *Journal of Cognitive Neuroscience*, 19(9), 1488-1497. doi:10.1162/jocn.2007.19.9.1488
- Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Behavioral and Brain Sciences*, 39. doi:10.1017/s0140525x15000965
- Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(1), 126-132. doi:10.1037//0278-7393.10.1.126
- Freyd, J.J., & Jones, K.T. (1994) Representational momentum for a spiral path. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 968-976. <http://dx.doi.org/10.1037/0278-7393.20.4.968>
- Frischen, A., & Tipper, S. P. (2006). Long-term gaze cueing effects: Evidence for retrieval of prior states of attention from memory. *Visual Cognition*, 14(3), 351-364. doi:10.1080/13506280544000192
- Friston, K., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks*, 22(8), 1093-1104. doi:10.1016/j.neunet.2009.07.023
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12), 493-501.
- Gao, T., McCarthy, G., & Scholl, B. J. (2010). The Wolfpack Effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science*, 21(12), 1845-1853. doi:10.1177/0956797610388814
- Gao, T., Scholl, B. J., & McCarthy, G. (2012). Dissociating the Detection of Intentionality from Animacy in the Right Posterior Superior Temporal Sulcus. *Journal of Neuroscience*, 32(41), 14276-14280. doi:10.1523/jneurosci.0562-12.2012
- Hammett, S. T. (1997). Motion blur and motion sharpening in the human visual system. *Vision Research*, 37(18), 2505-2510. doi:10.1016/s0042-6989(97)00059-x
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243–259. <https://doi.org/10.2307/1416950>
- Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, 12(5), 822-851. doi:10.3758/bf03196775
- Hubbard, T. L. (2006). Bridging the gap: possible roles and contributions of representational momentum. *Psicológica*, 27(1), 1–34.
- Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics*, 44(3), 211-221.
- Hudson, M., Bach, P., & Nicholson, T. (2017). You said you would! The predictability of other’s behavior from their intentions determines predictive biases in action perception. *Journal of Experimental Psychology: Human Perception and Performance*, 44(2), 320-335. doi:10.1037/xhp0000451

- Hudson, M., McDonough, K. L., Edwards, R., & Bach, P. (2018). Perceptual Teleology: Expectations of Action Efficiency Bias Social Perception. *Proceedings of the Royal Society: B*, (285(1884)), 20180638. doi:10.1098/rspb.2018.0638
- Hudson, M., Nicholson, T., Ellis, R., & Bach, P. (2016). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions. *Cognition*, 146, 245-250. doi:10.1016/j.cognition.2015.09.021
- Hudson, M., Nicholson, T., Simpson, W. A., Ellis, R., & Bach, P. (2016). One step ahead: The perceived kinematics of others' actions are biased toward expected goals. *Journal of Experimental Psychology: General*, 145(1), 1-7. doi:10.1037/xge0000126
- Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, 9(1), 21-25.
- Jacquet, P. O., Chambon, V., Borghi, A. M., and Tessari, A. (2012). Object affordances tune observers' prior expectations about tool-use behaviors. *PLoS One* 7:e39629. doi: 10.1371/journal.pone.0039629
- Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural Representation for the Perception of the Intentionality of Actions. *Brain and Cognition*, 44(2), 280-302. doi:10.1006/brcg.2000.1231
- Joyce, K., Schenke, K., Bayliss, A., & Bach, P. (2015). Looking ahead: Anticipatory cueing of attention to objects others will look at. *Cognitive Neuroscience*, 7(1-4), 74-81. doi:10.1080/17588928.2015.1053443
- Kalénine, S., Shapiro, A. D., Flumini, A., Borghi, A. M., & Buxbaum, L. J. (2014). Visual context modulates potentiation of grasp types during semantic object categorization. *Psychonomic Bulletin & Review*, 21(3), 645-651.
- Kerzel, D. (2003). Mental extrapolation of target position is strongest with weak motion signals and motor responses. *Vision Research*, 43(25), 2623-2635.
- Keulen, R. F., Adam, J. J., Fischer, M. H., Kuipers, H., & Jolles, J. (2003). Distractor interference in selective reaching: Dissociating distance and grouping effects. *Journal of motor behavior*, 35(2), 119-126.
- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007a). The mirror-neuron system: a Bayesian perspective. *NeuroReport*, 18(6), 619-623. doi:10.1097/wnr.0b013e3281139ed0
- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007b). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, 8(3), 159-166. doi:10.1007/s10339-007-0170-2
- Kok, P., Jehee, J. F., & De Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, 75(2), 265-270.
- Koster-Hale, J., & Saxe, R. (2013). Theory of mind: a neural prediction problem. *Neuron*, 79(5), 836-48. doi:10.1016/j.neuron.2013.08.020
- Kozhevnikov, M., & Hegarty, M. (2001). Impetus beliefs as default heuristics: Dissociation between explicit and implicit knowledge about motion. *Psychonomic Bulletin & Review*, 8(3), 439-453.

- Lakens, D. (2017). Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355-362. doi: 10.1177/1948550617697177
- Lakens, D., & Albers, C. (2017). When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of Experimental Social Psychology*, 74, 187-195.
- Lamme, V. A. F., Zipser, K., & Spekreijse, H. (2002). Masking Interrupts Figure-Ground Signals in V1. *Journal of Cognitive Neuroscience*, 14(7), 1044-1053. doi:10.1162/089892902320474490
- McDonough, K. L., Hudson, M., & Bach, P. (2019). Cues to intention bias action perception toward the most efficient trajectory. *Scientific Reports*, 9.
- Motes, M. A., Hubbard, T. L., Courtney, J. R., & Rypma, B. (2008). A principal components analysis of dynamic spatial memory biases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(5), 1076.
- Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary Visual Cortex Activity along the Apparent-Motion Trace Reflects Illusory Perception. *PLoS Biology*, 3(8), e265. doi:10.1371/journal.pbio.0030265
- Müsseler, J., Stork, S., & Kerzel, D. (2008). Localizing the onset of moving stimuli by pointing or relative judgment: Variations in the size of the Fröhlich effect. *Vision Research*, 48(4), 611-617.
- Nicholson, T., Roser, M., & Bach, P. (2017). Understanding the Goals of Everyday Instrumental Actions Is Primarily Linked to Object, Not Motor-Kinematic, Information: Evidence from fMRI. *PLOS ONE*, 12(1), e0169700. doi:10.1371/journal.pone.0169700
- Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, 370(6487), 256-257. doi:10.1038/370256b0
- Nijhawan, R. (2008). Visual prediction: Psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences*, 31(02). doi:10.1017/s0140525x08003804
- Pellicano, E., & Burr, D. (2012). When the world becomes “too real”: A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16(10), 504–510. doi:10.1016/j.tics.2012.08.009
- Prinz, W., & Rapinett, G. (2008). Filling the gap: Dynamic representation of occluded action. *Emerging Communication*, 10(223).
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nat Rev Neuro*, 11, 264-274. doi: 10.1038/nrn2805.
- Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., & Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiological Reviews*, 94(2), 655-706.

- Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition*, 119(2), 242-252. doi:10.1016/j.cognition.2011.01.014
- Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, 42(11), 1435-1446. doi:10.1016/j.neuropsychologia.2004.04.015
- Schäfer, T., & Schwarz, M. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, 10, 813.
- Schenke, K. C., Wyer, N. A., & Bach, P. (2016). The Things You Do: Internal Models of Others' Expected Behaviour Guide Action Observation. *PLOS ONE*, 11(7), e0158910. doi:10.1371/journal.pone.0158910
- Scholl, B. J., & Gao, T. (2013). Perceiving Animacy and Intentionality: Visual processing or higher-level judgment. In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social Perception: Detection and interpretation of animacy, agency, and intention*, (pp. 197-230): The MIT Press.
- Southgate, V., Johnson, M. H., Karoui, I. E., & Csibra, G. (2010). Motor System Activation Reveals Infants' On-Line Prediction of Others' Goals. *Psychological Science*, 21(3), 355-359. doi:10.1177/0956797610362058
- Uithol, S., van Rooij, I., Bekkering, H., & Haselager, P. (2011). What do mirror neurons mirror? *Philosophical Psychology*, 24(5), 607-623.
- van Elk, M., van Schie, H., & Bekkering, H. (2014). Action semantics: A unifying conceptual framework for the selective use of multimodal and modality-specific object knowledge. *Physics of Life Reviews*, 11(2), 220-250. doi:10.1016/j.plrev.2013.11.005
- von der Lüh, T., Manera, V., Barisic, I., Becchio, C., Vogeley, K., & Schilbach, L. (2016). Interpersonal predictive coding, not action perception, is impaired in autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693), 20150373. doi:10.1098/rstb.2015.0373
- Welsh, T. N., Elliott, D., & Weeks, D. J. (1999). Hand deviations toward distractors Evidence for response competition. *Experimental Brain Research*, 127(2), 207-212.
- Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications*, 9(1), 4288.

Footnotes

¹Please note that although our study design is based on prior work on Representational Momentum (e.g. Hubbard, 2005; Hubbard & Bharucha, 1988), the predicted effects are not along the trajectory of motion (the X-axis), typically referred to as "M-Displacement". Instead, perceptual biases are predicted along the axis perpendicular to the direction of motion and towards the apparent target object (the Y-axis). These displacements are typically referred to as "O-Displacement". We first calculate perceptual biases as differences in screen coordinates between the actual stimulus location and participant's touch-screen responses, making both the apparent shifts towards the relevant goal object (along the Y-axis) and the general shift towards the hands' centre of gravity (along the X-axis) directly visible. We then provide a summary measure of the predicted perceptual displacements on the Y-axis ("O-Displacement"), reported in Figure 2C.